PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/31, 15/47, 15/62, 15/74, 1/21, C07K 14/145, 14/21

A1

(11) International Publication Number: WO 97/34000

(43) International Publication Date: 18 September 1997 (18.09.97)

(21) International Application Number: PCT/CA97/00167

(22) International Filing Date: 10 March 1997 (10.03.97)

(30) Priority Data:
08/614,377 12 March 1996 (12.03.96) US

(60) Parent Applications or Grants
(63) Related by Continuation
US 07/895,367 (CIP)
Filed on 9 June 1992 (09.06.92)
US 08/194,290 (CIP)
Filed on 9 February 1994 (09.02.94)

(71) Applicant *(for all designated States except US)*: THE UNIVERSITY OF BRITISH COLUMBIA [CA/CA]; IRC Room 331, 2194 Health Sciences Mall, Vancouver, British Columbia V6T 1Z3 (CA).

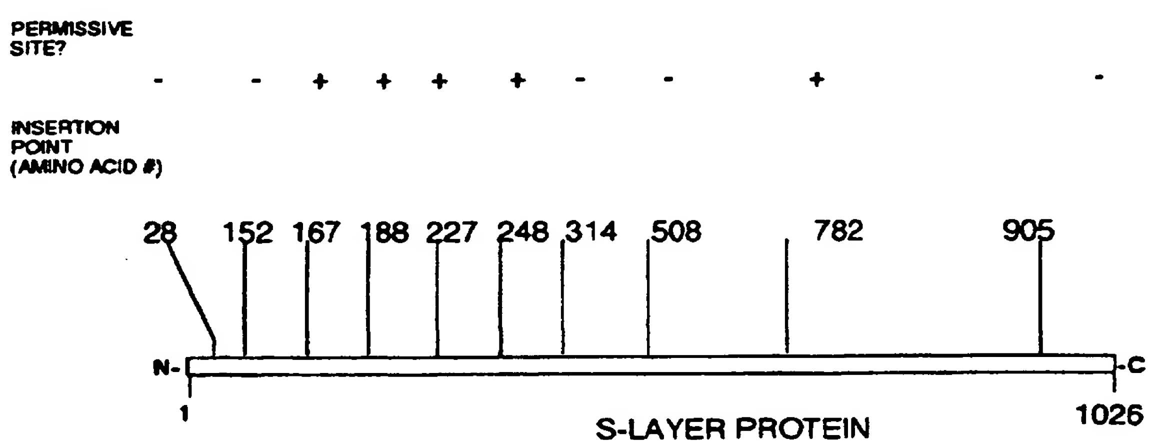
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: SMIT, John [US/CA]; 9960 Seacastle Drive, Richmond, British Columbia V7A 4R8 (CA). BINGLE, Wade, H. [CA/CA]; 111 - 1370 Davie Street, Vancouver, British Columbia V6E 1N6 (CA). NOMELLINI, John, F. [US/CA]; 6451 Riverdale Drive, Richmond, British Columbia V7C 2G2 (CA).

(74) Agents: ROBINSON, J., Christopher et al.; Smart & Biggar, Box 11560, Vancouver Centre, Suite 2200, 650 West Georgia Street, Vancouver, British Columbia V6B 4N8 (CA).

(81) Designated States: AU, CA, IL, JP, SG, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: EXPRESSION AND SECRETION OF HETEROLOGOUS POLYPEPTIDES FROM CAULOBACTER



(57) Abstract

DNA constructs are provided which code for a chimeric protein in which the C-terminal region correponds to the extreme C-terminal amino acids of a *Caulobacter* S-layer protein, fused with a heterologous polypeptide. Bacterial cells containing, or which express the DNA c nstructs and secrete the resulting protein are also provided. Chimeric proteins including the C-terminal amino acids of a *Caulobacter* S-layer protein are provided, including proteins which include antigenic epitopes of the Infectious Hematopoietic Necrosis Virus.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

Expression and Secretion of Heterologous Polypeptides from Caulobacter

Field of Invention

This invention relates to the expression and secretion of heterologous peptides, from Caulobacter wherein the heterologous polypeptide is fused with the surface layer protein (S-layer protein) of the bacterium, or a portion of the S-layer protein.

Background of the Invention

Bacterial surface proteins have been used as carriers for foreign (heterologous) polypeptides (particularly in Salmonella and E. coli) for various purposes, including the development of live vaccines. In some instances, the heterologous material is expressed as a fusion product with a surface protein of the bacterium. Generally, the use of such surface proteins as a vehicle for expression and/or presentation of heterologous polypeptides has been limited by the characteristics of a particular surface protein. The lipopolysaccharide layer of a bacterium, which tends to stimulate a strong immune response, covers the integral outer membrane proteins of the organism and potentially affects efficient presentation of a cloned epitope. Where the surface protein is functional (for example, as part of a filamentous portion of a bacterial cell surface) there will be limited opportunities to express a fusion product and still retain the surface protein's function. Generally, the organisms that have been used for these purposes have been chosen because of the advantages presented in respect of the organism's relationship to its host.

Many genera of bacteria assemble layers composed of repetitive, regularly aligned, proteinaceous sub-units on the outer surface of the cell. These layers are essentially two-dimensional paracrystalline arrays, and

being the outer molecular layer of the organism, directly interface with the environment. Such layers are commonly known as S-layers and are found on members of every taxonomic group of walled bacteria including: Archaebacteria; Chlamydia; Cyanobacteria; Acinetobacter; Bacillus; Aquaspirillum; Caulobacter; Clostridium; Chromatium. Typically, an S-layer will be composed of an intricate, geometric array of at least one major protein having a repetitive regular structure. In many cases, such as in Caulobacter, the S-layer protein is synthesized by the cell in large quantities and the S-layer completely envelopes the cell and thus appears to be a protective layer.

Caulobacter are natural inhabitants of most soil and freshwater environments and may persist in waste water treatment systems and effluents. The bacteria alternate between a stalked cell that is attached to a surface, and an adhesive motile dispersal cell that searches to find a new surface upon which to stick and convert to a stalked cell. The bacteria attach tenaciously to nearly all surfaces and do so without producing the extracelluar enzymes or polysaccharide "slimes" that are characteristic of most other surface attached bacteria. They have simple requirements for growth. The organism is ubiquitous in the environment and has been isolated from oligotrophic to mesotrophic situations. Caulobacters are known for their ability to tolerate low nutrient level stresses, for example, low phosphate levels. This nutrient can be limiting in many leachate waste streams, especially those with high levels of iron or calcium.

All of the freshwater Caulobacter that produce an S-layer are similar and have S-layers that are substantially the same. Such S-layers appear similar by electron microscopy with the layer being hexagonally arranged in all cases with a similar centre-centre dimension (see: Walker,

S.G., et al. (1992). "Isolation and Comparison of the Paracrystalline Surface Layer Proteins of Freshwater Caulobacters" J. Bacteriol. 174: 1783-1792). 16S rRNA sequence analysis of several S-layer producing Caulobacter strains suggest that they group closely (see: Stahl, D.A. et al (1992) "The Phylogeny of Marine and Freshwater Caulobacters Reflects Their Habitat" J. Bacteriol. 174:2193-2198). DNA probing of Southern blots using the S-layer gene from C. crescentus CB15 identifies a single band that is consistent with the presence of a cognate gene (see: MacRae, J.D. and, J. Smit. (1991) "Characterization of Caulobacters Isolated from Wastewater Treatment Systems" Applied and Environmental Microbiology 57:751-758). Furthermore, antisera raised against the S-layer protein of C. crescentus strain CB15 reacts with S-layer proteins from other Caulobacter (see: Walker, S.G. et al (1992) [supra]). All S-layer proteins isolated from Caulobacter may be substantially purified using the same extraction method (pH extraction) which would not be expected to be a general purpose method for other bacterial membrane or surface associated proteins. All strains appear to have a polysaccharide reactive with antisera reactive against CB15 lipopolysaccharide species which appears to be required for S-layer attachment (see: Walker, S.G. et al (1992) [supra]).

The S-layer elaborated by freshwater isolates of Caulobacter are visibly indistinguishable from the S-layer produced by Caulobacter crescentus strains CB2 and CB15. The S-layer proteins from the latter strains have approximately 100,000 m.w. although sizes of S-layer proteins from other species and strains will vary. The protein has been characterized both structurally and chemically. It is composed of ring-like structures spaced at 22nm intervals arranged in a hexagonal manner on the outer membrane. The S-layer is bound to the bacterial

surface and may be removed by low pH treatment or by treatment with a calcium chelator such as EDTA.

The similarity of S-layer proteins in different strains of Caulobacter permits the use of a cloned S-layer protein gene of one Caulobacter strain for retrieval of the corresponding gene in other Caulobacter strains (see: Walker, S.G. et al (1992) [supra]; and, MacRae, J.D. et al (1991) [supra].

Expression, secretion and optionally, presentation, of a heterologous polypeptide as a fusion product with the S-layer protein of Caulobacter provides advantages not previously seen in systems using organisms such as E. coli and Salmonella where fusion products of other kinds of surface proteins have been expressed. All known Caulobacter strains are believed to be harmless and are nearly ubiquitous in aquatic environments. In contrast, many Salmonella and E. coli strains are pathogens. Consequently, expression and secretion of a heterologous polypeptide using Caulobacter as a vehicle will have the advantage that the expression system will be stable in a variety of outdoor environments and may not present problems associated with the use of a pathogenic organism. Furthermore, Caulobacter are natural biofilm forming species and may be adapted for use in fixed biofilm bioreactors. The quantity of S-layer protein that is synthesized and is secreted by Caulobacter is high, reaching 12% of the cell protein. The unique characteristics of the repetitive, two-dimensional S-layer would also make such bacteria ideal for use as an expression system, or as a presentation surface for heterologous polypeptides. This is desirable in a live vaccine to maximize presentation of the antigen or antigenic epitope. In addition, use of such a presentation surface to achieve maximal exposure of a desired polypeptide to the environment results in such bacteria

being particularly suited for use in bioreactors or as carriers for the polypeptide in aqueous or terrestrial outdoor environments.

Summary of Invention

This invention pertains to the discovery that the C-terminal region of the Caulobacter S-layer protein is essential for secretion of the S-layer protein. The inventors have determined that the 3' region of the gene which encodes the C-terminal region of the S-layer protein is conserved among different strains of Caulobacter.

This invention provides a method of expressing and presenting to the environment of a Caulobacter, a polypeptide that is heterologous to the S-layer protein of the Caulobacter, which comprises inserting a coding sequence for the heterologous polypeptide in-frame into a S-layer protein gene of Caulobacter, or a portion of said S-layer protein gene, whereby the polypeptide is expressed and secreted by the Caulobacter as a chimeric protein comprising the heterologous protein and all or part of the S-layer protein.

This invention provides a DNA construct for the aforemention chimeric protein, and a bacterium comprising such a DNA construct, wherein the DNA construct encodes all or part of a S-layer protein, and one or more in-frame sequences encoding one or more heterologous proteins.

This invention provides a DNA construct comprising one or more restriction sites for facilitating insertion of DNA into the construct and, DNA encoding at least the 82 C-terminal amino acids of Caulobacter S-layer protein. Preferably, the C-terminal amino acids are or correspond to amino acids 944 or 945-1026 of the RsaA protein of C. crescentus.

This invention provides a DNA construct comprising DNA encoding a heterologous polypeptide sequence not present in a Caulobacter S-layer protein upstream from and in-frame with DNA encoding at least the 82 C-terminal amino acids of Caulobacter S-layer protein. Preferably, the C-terminal amino acids are or correspond to amino acids 944 or 945-1026 of the rsaA protein of C. crescentus.

This invention also provides a secreted protein obtained from the cell surface or cell medium of a Caulobacter cell expressing the aforementioned DNA constructs wherein the secreted protein comprises the heterologous polypeptide and at least the 82 C-terminal amino acids of a Caulobacter S-layer protein. Preferably, the C-terminal amino acids are or correspond to amino acids 944 or 945-1026 of the RsaA protein of C. crescentus.

Description of the Drawings

For better understanding of this invention, reference may be made to the preferred embodiments and examples described below, and the accompanying drawings in which:

Figure 1 is the sequence of a Carrier cassette which may be cloned into the PstI/BamHI site of pUC9 to deliver a gene sequence of interest to sites within a Caulobacter crescentus S-layer protein (rsaA) gene (SEQ ID NO:1).

Figure 2 is a restriction map of a plasmid based promoter-less version of the rsaA gene (pTZ18U:rsaAΔP) containing restriction sites and which may be used to accept heterologous DNA of interest.

Figure 3 is the nucleotide sequence of linker BamHI-7165K (SEQ ID NO:2; and SEQ ID NO:3) carried in plasmid pUC9B (pUC7165K), which may be used for mutagenesis at

sites created in rsaA by a specific or non-specific endonuclease.

Figure 4 is the nucleotide sequence a linker BamHI-6571K (SEQ ID NO:4; and SEQ ID NO:5) carried in plasmid pTZ19 (pTZ6571K) which may be used for mutagenesis at sites created in rsaA by a specific or non-specific endonuclease.

Figure 5 is a map of insertion events at TaqI sites in the rsaA gene identified by amino acid number of the insertion site in the S-layer protein and scored according to whether the S-layer is produced in the modified organism.

Figure 6 (comprising figures 6a, b, and c) shows the complete nucleotide sequence of the C. crescentus S-layer (rsaA) gene (SEQ ID NO:6) and the predicted translational product in the single letter amino acid code. The -35 and -10 sites of the promoter region as well as the start of transcription and the Shine-Dalgarno sequence are indicated. Partial amino acid sequences determined by Edman degradation of rsaA protein and of sequenced peptides obtained after cleavage with V8 protease are indicated by contiguous underlining. The putative transcription terminator palindrome is indicated with arrowed lines. The region encoding the glycine-aspartate repeats is indicated by underlined amino acid code letters. This region includes five aspartic acids that may be involved in the binding of calcium ions.

Figure 7 is a bar graph showing the approximate location by amino acid block of 54 permissive sites in the rsaA gene corresponding to TaqI, HinPI, AciI, and MspI sites described in Example 3.

Figure 8 is a portion of an amino acid sequence (SEQ ID NO: 8) from P. aeruginosa PAK pilin in which the 12

amino acid pilus peptide epitope referred to in Example 5 is identified by superscript numerals 1-12.

Figure 9 is the nucleotide coding sequence and corresponding amino acid sequence (SEQ ID NO:9) in respect of the 184 amino acid sequence corresponding to amino acids 270-453 of the IHNV surface glycoprotein described in Example 6.

Figure 10 is the amino acid sequence of the synthetic cadmium binding peptide referred to in Example 4. The cadmium binding site is shown in the figure.

Figure 11 shows locations of some of the sites in _rsa_A in which single and multiple copies of the pilus peptide described in Example 5 was expressed and secreted as part of a chimeric _rsa_A protein.

Figure 12 shows a portion of pUC8 containing various C-terminal fragments of _rsa_A as described in Example 7.

Description of the Preferred Embodiments

The preferred organism for use in this invention is _Caulobacter_, particularly _C. crescentus_. While similarity of the S-layer gene and S-layer secretion systems permits the use of any S-layer protein producing _Caulobacter_ in this invention, _C. crescentus_ strains CB2 and CB15 and variants of those strains which contain homologs of the gene encoding the 1026 amino acid paracrystalline S-layer protein described in: Gilchrist, A. _et al_. 1992. "Nucleotide Sequence Analysis Of The Gene Encoding the _Caulobacter crescentus_ Paracrystalline Surface Layer Protein". Can. J. Microbiol. 38:193-208, are referred to in the examples described below.

Caulobacter strains which either are incapable of forming an S-layer, including those which shed the S-layer

protein upon secretion, may be used in this invention. Examples are the variants CB2A and CB15AKSac described in Smit, J., and N. Agabian. 1984. "Cloning of the Major Protein of the Caulobacter crescentus Periodic Surface Layer: Detection and Other Characterization of the Cloned Peptide by Protein Expression Assays". J. Bacteriol. 160:1137-1145.; and, Edwards, P., and J. Smit. 1991. "A Transducing Bacteriophage for Caulobacter crescentus Uses the Paracrystalline Surface Layer Protein as Receptor". J. Bacteriol. 173, 5568-5572. Examples of shedding strains are CB15Ca5 and CB15Ca10 described in Edwards and Smit (1991) [supra], and the smooth lipopolysaccharide deficient mutants described in Walker, S.G. et al. 1994. "Characterization of Mutants of Caulobacter crescentus Defective in Surface Attachment of the Paracrystalline Surface Layer". J. Bacteriol. 176:6312-6323.

A heterologous polypeptide referred to herein may be any peptide, polypeptide, protein or a part of a protein which is desired to be expressed in Caulobacter and which may be secreted by the bacterium. The heterologous polypeptide includes enzymes and other functional sequences of amino acids as well as ligands, antigens, antigenic epitopes and haptens. The size of the heterologous polypeptide will be selected depending upon whether an intact S-layer is to be produced in the Caulobacter or whether the chimeric protein to be recovered from the bacterial medium as described below. Preferably, the cysteine content of the heterologous polypeptide and the capacity for formation of disulphide bonds within the chimeric protein will be kept to a minimum to minimize disruption of the secretion of the chimeric protein. However, the presence of cysteine residues capable of forming a disulphide bond which are relatively close together, may not affect secretion.

Once a particular bacterium's S-layer protein gene is characterized, this invention may be practised by implementing one or more known methods to insert a selected heterologous coding sequence into all or part of the S-layer protein gene so that both the S-layer protein and the heterologous sequence are transcribed "in-frame". Knowledge of an S-layer protein gene sequence permits one to identify potential sites to install the heterologous genetic material. The repetitive nature of the protein in the S-layer permits multiple copies of a heterologous polypeptide to be presented on the surface of the cell.

The following general procedure lays out courses of action and specifies particular plasmid vectors or constructions that may be used to accomplish fusion of an S-Layer protein with a polypeptide of interest. The following description uses the rsaA (S-layer) gene of C. crescentus as an example (see Figure 6 and SEQ ID NO:6). The latter gene sequence is characterized in Gilchrist, A. et al (1992) [supra).

The general procedure includes detailed steps allowing for the following possibilities:

1) use of a collection of potentially permissive sites in the S-layer gene to install the genetic information for a polypeptide of interest;

2) use of a Carrier cassette for delivering a gene of interest to sites within the S-layer gene (the cassette offers several advantages over direct modification of a gene of interest, in preparation for insertion);

3) creation of a collection of random insertion sites based on a restriction enzyme of choice, if the available collection of potentially permissive sites is for some reason unsuitable; and,

4) preparation of DNA coding for a polypeptide of interest for direct insertion into permissive sites (ie,

not using the Carrier cassette) by a method best suited for the particular case (several options are suggested).

The general procedure involves the following steps and alternative courses of action. As a first step the practitioner will choose an appropriate region (or specific amino acid position) of the S-layer for insertion of a desired polypeptide. Second, the practitioner will create a unique restriction site (preferably hexameric) in the rsaA (S-layer) gene at a position within the gene encoding that region (or corresponding to a specific amino acid) using either standard linker mutagenesis (regional) or site directed mutagenesis (specific amino acid). The unique restriction site will act as a site for accepting DNA encoding the polypeptide of interest. The plasmid-based promoter-less version of the rsaA gene (pTZ18U:rsaAΔP) shown in Figure 2 may be used because it contains an appropriate combination of 5′ and 3′ restriction sites useful for subsequent steps (see: Gilchrist, A. et al (1992) [supra]). The restriction site should not occur in rsaA, its carrier plasmid or the DNA sequence coding for the polypeptide of interest.

If it is unclear which region of the S-layer would be suitable for insertion of a polypeptide of interest, a random linker mutagenesis approach is used to randomly insert a unique linker-encoded restriction site (preferably hexameric) at various positions in the rsaA gene. Sites for insertion of the linker are created using an endonuclease, either of a sequence specific nature (e.g. tetrameric recognition site restriction enzyme) or sequence non-specific nature (e.g. Deoxyribonuclease I [DNase I]). A particularly suitable method is the generalized selectable linker mutagenesis approach based on any desired restriction site of: Bingle, W.H., and J. Smit. 1991 "Linker Mutagenesis Using a Selectable Marker: A Method for Tagging Specific Purpose Linkers With an Antibiotic-

Resistance Gene". Biotechniques 10: 150-152. Because endonuclease digestion is carried out under partial digestion conditions, a library of linker insertions at different positions in rsaA is created. Partial digestion with MspI, HinPI and Aci:I can create 150 potential sites for insertion of a Bam HI linker such as:

5'-CGACGGATCCGT
TGCCTAGGCAGC-5'
(SEQ ID. NO:10).

If restriction endonucleases are used to create sites for subsequent insertion of a linker encoding a hexameric restriction site, mutagenesis may also be done with a mixture of 3 different linkers incorporating appropriate spacer nucleotides in order to satisfy reading frame considerations at a particular restriction site (only 1 of the 3 linker insertions will be useful for subsequent acceptance of DNA encoding the polypeptide of interest). With DNase I, only one linker is needed, but again only 1 of 3 linker insertions may be useful for accepting DNA encoding the polypeptide of interest depending on the position of the DNase I cleavage with respect to the 3 bases of each amino acid codon.

Next, a linker tagged with a marker is used to insert DNA of interest at a restriction site. For example, if BamHI sites are appropriate as sites for the introduction of DNA encoding a polypeptide of interest, BamHI linkers tagged with a kanamycin-resistance gene for selectable linker mutagenesis may be used. One such 12-bp linker carried in plasmid pUC1021K was described by Bingle and Smit (1991) [Supra]. Two additional 15-bp linkers (pUC7165K and pTZ6571K) constructed for creating the other 2 possible translation frames within the linker insert itself are described in Figures 3 and 4 (SEQ ID NO:2; SEQ ID NO:3; SEQ ID NO:4; and, SEQ ID NO:5). Any one of the above three kanamycin-resistance tagged BamHI linkers is

suitable for mutagenesis at sites created in rsaA by DNase I. As outlined above, a mixture of all three linkers is preferably used for mutagenesis at sites created in rsaA by restriction enzyme digestion.

Once a library composed of linker insertions encoding desired hexameric restriction site at different positions in rsaA has been created, the DNA encoding a polypeptide of interest is inserted into the sites en masse (the library of mutated rsaA genes may be manipulated as one unit). The library is digested with the restriction enzyme specific for the newly-introduced linker encoded restriction site and ligated to a DNA fragment encoding the polypeptide of interest and carrying the appropriate complementary cohesive termini. The DNA specifying the polypeptide of interest can be prepared by a number of standard methods, which may include oligonucleotide synthesis of 2 anti-complementary strands, polymerase chain reaction (PCR) procedures, or addition of linkers whose termini are compatible with the introduced sites in rsaA to a suitably modified segment of DNA.

In order to facilitate the rapid recovery of useful rsaA genes carrying newly inserted DNA at BamHI sites encoding the polypeptide of interest, the Carrier oligonucleotide shown in Figure 1 may be used. The Carrier is designed to accept DNA (including multiple copies and mixtures) prepared by PCR or annealed synthesized oligonucleotides and controls direction of insertion of the foreign segment into a rsaA gene through use of a promoterless drug resistance marker. The DNA of interest is first directionally cloned, if possible, using the XhoI, StuI, or SalI sites or non-directionally cloned using any one of the sites in the same orientation as a promoterless chloramphenicol resistance (CmR) gene. To do this the DNA of interest must be provided with the appropriate termini for cloning and spacer nucleotides for maintaining correct

reading frame within the cassette and should not contain a BglII site. For insertion into the BamHI linker library, the DNA of interest is recovered as a BamHI fragment tagged with a CmR gene. When ligated to the BamHI digested rsaA linker library, only those colonies of the bacterium (eg. E. coli) used for the gene modification steps that are recovered will be those carrying insertions of the desired DNA in the correct orientation, since the promoter on the plasmid is 5′ to rsaAΔP and the CmR gene. This eliminates screening for DNA introduction and increases the recovery of useful clones by 100% (1 of 3 versus 1 of 6). While still manipulating the library as one unit, the CmR gene is removed using BglII. The carrier oligonucleotide also provides the opportunity to add DNA 5′ or 3′ to the DNA of interest at SalI, XhoI or StuI sites providing the DNA of interest does not contain any of these sites. This allows some control over spacing between rsaA sequences and the sequence of the DNA of interest.

Next, the rsaA genes carrying the DNA of interest in the correct orientation is excised from the plasmid (eg. from the pTZ18U:rsaAΔP plasmid) and is transferred to a suitable vector providing a promoter recognized by Caulobacter. Such vectors include pWB9 or pWB10 (as described in Bingle, W.H., and J. Smit. 1990). "High Level Plasmid Expression Vectors for Caulobacter crescentus Incorporating the Transcription and Transcription-Translation Initiation Regions of the Paracrystalline Surface Layer Protein Gene". Plasmid 24: 143-148) with EcoRI/SstI sites. The DNA of interest should not contain the same restriction sites present in the vector. The latter vectors allow expression of rsaA hybrids in S-layer negative mutants of Caulobacter such as CB15KASac.

Those Caulobacter surviving transfer are examined for chimeric protein secretion, S-layer assembly and presentation of the new polypeptide activity, antigenicity,

etc. by methods specific to the needs of the investigator or the capabilities of the inserted sequence. Many of the sites created are "benign" as they have no effect on the functional regions of the protein involved with export, self assembly, etc. However, not every site that results in an absence of functional disruption of the S-layer is best for insertion of new activities. Some sites may not be well exposed on the surface of the organism and other sites may not tolerate insertion of much more DNA than the linker sequence.

By selecting the site of insertion of the heterologous material, it is possible to express heterologous polypeptides of up to about 60 (preferably less than 50) amino acids in a S-layer chimeric protein which will assemble as an S-layer on the cell surface. Single or multiple insertions of smaller polypeptides (eg. 10-20 amino acids) at a wide range of the permissive sites in the S-layer gene will permit S-layer formation. Some sites, as reported herein, are sensitive to even small insertions resulting in the chimeric protein being released into the medium. Release may also be deliberately effected by use of a shedding strain of Caulobacter to express the chimeric protein or by physical removal of the S-layer from whole cells.

Where S-layer formation is not required, this invention permits the expression of quite large polypeptides (eg. about 200 amino acids) as part of the S-layer protein. Expressing a chimeric protein containing a S-layer protein component having substantial deletions, as described below, may increase the size of the heterologous polypeptides that will be expressed and secreted by Caulobacter.

The preceding methods describe insertion of linkers in-frame into an rsaA gene (eg. a promoterless version of

the gene). The sites that are introduced allow subsequent insertion of foreign DNA in-frame into the full length rsaA gene. This invention also includes the construction of chimeric S-layer protein genes and the resulting production of chimeric S-layer proteins wherein the S-layer gene component is highly modified by deleting large portions of that gene which reduces the amount of Caulobacter protein present in the secreted chimeric protein.

Generally, large deletions throughout the S-layer gene will result in a chimeric protein that is not capable of forming an S-layer. Attachment of the S-layer to the cell is abolished if about the first 29 N-terminal amino acids of the S-layer protein are deleted. Deletion of the first 776 amino acids from the N-terminal region will still result in a chimeric protein that is secreted from the cell but having a S-layer protein component of only the 250 C-terminal amino acids. It has also been found that only the extreme C-terminal region corresponding to approximately amino acids 945-1026 of RsaA is required for secretion of an S-layer chimeric protein from Caulobacter. Thus the chimeric protein need only have the 82 amino acid C-terminal region of the S-layer protein to be secreted from the cell. Furthermore, use of the C-terminal region corresponding to about amino acids 850-1026 (or more) of RsaA not only permits the cell to transport the chimeric protein outside of the cell, but also promotes spontaneous aggregation of much of the secreted chimeric protein in the cell medium and formation of a macroscopic precipitate that may be collected with a course mesh or sheared to micron-sized particles which may be ideal for vaccine presentation. Yields of up to 250 mg. (dry weight) of protein per liter of cells may be possible.

Sequence analysis of the 3' region of the S-layer genes from different strains of Caulobacter shows that the portion of the gene encoding the C-terminal region of the

S-layer protein is highly conserved along with the immediate downstream non-translated and translated region. Sequence analysis of the S-layer genes and downstream regions in CB15 and CB2A (which are readily distinguishable strains) shows identical DNA sequences coding for the last 118 C-terminal amino acids of the S-layer protein and the downstream non-translated region. Sequencing of the next downstream translated gene to amino acid 97 of the gene product shows only a single base pair change between CB15 and CB2A, resulting in a conservative amino acid substitution in the translation product. Conservation of the C-terminal region of Caulobacter S-layer protein and associated coding regions shows that this invention may be carried out using any Caulobacter producing a S-layer protein.

This invention may be practised as shown in the Examples by expression of modified S-layer genes borne on plasmids that are broad host range vectors capable of being expressed in Caulobacter. Such plasmids are readily constructed and introduced to Caulobacter by electroportation. Typically, the plasmid is maintained in the Caulobacter by antibiotic selection. Highly modified rsaA genes with attached heterologous sequences may also be introduced into Caulobacter on a plasmid that is not replicated by Caulobacter. At a low but practicable frequency, homologous recombination of the incoming modified S-layer gene with the chromosome-resident copy of the S-layer gene in the cell will result in a gene rescue or transfer event. In some cases it may be desirable to obtain a stable cell line in which the chimeric S-layer gene is chromosomal. Various protocols for creating chromosomal insertions are set out in the Examples.

Use of the S-layer protein as a vehicle for production of a heterologous polypeptide has several advantages. Firstly, the S-layer protein is synthesized in large

quantities and has a generally repetitive sequence. This permits the development of systems for synthesis of a relatively large amount of heterologous material as a fusion product with an S-layer protein (chimeric protein). It may be desirable to retain the chimeric protein as part of the bacterial cell envelope or, the fusion product may be separated from the organism, such as by the method described in: Walker, S.G., et al. 1992. "Isolation and Comparison of the Paracrystalline Surface Layer Proteins of Freshwater Caulobacters". J. Bacteriol. 174:1783-1792. Alternatively, the Caulobacter strain that is used to express the fusion product may be derived from a strain such as CB15Ca5 that sheds its S-layer.

This invention is particularly suited for use in a bioreactor systems. An example would be the use of a modified Caulobacter expressing a polypeptide having activity similar to that of a metallothionein in a bioreactor, to bind toxic metals in sewage, waste water etc. Caulobacters are ideal candidates for fixed-cell bioreactors, the construction of which is well known. An example of such a bioreactor is a rotating biological contactor. Although other bacteria are found in the environment that are capable of binding metals, they often do so by producing copious polysaccharide slimes that quickly plug filtration systems. In some cases, the bacteria are not surface-adherent or the bacteria do not show selectivity towards key toxic metals. By taking advantage of the natural bio-film forming characteristics of Caulobacter, bioreactors may be formed comprising a substrate and a single layer of cells adhered thereon, with the cells distributed at high density. A variety of substrates may be used such as a column of chemically derivatized glass beads or a porous ceramic material such as ceramic foam.

Another advantageous application for this invention is in the production of batch cultures of modified Caulobacter wherein the S-layer protein is a fusion product with an enzyme. For example, such Caulobacter could be grown in wood pulp suspensions at an appropriate juncture of the pulping process in order to provide for enzymatic decomposition of the wood-pulp structure (e.g. with an enzyme having an activity like xylanase or cellulase). Such an application may permit more effective penetration of bleaching agents in the wood-pulp bleaching process thereby reducing the use of chlorine-based bleaching agents.

Examples of enzymes that may be expressed as chimeric S-layer proteins include alkaline phosphatase (eg. by expression of the pho A gene of E. coli; see: Hoffman, C.S., and Wright, A. 1985. "Fusions of Secreted Protein to Alkaline Phosphatase: An Approach for Studying Protein Secretion". Proc. Natl. Acad. Sci. U.S.A. 82:5107-5111; Bingle, W.H., et al. 1993." An "All Purpose" Cellulase Reporter for Gene Fusion Studies and Application to the Paracrystalline Surface (S)-Layer Protein of Caulobacter crescentus". Can.J. Microbiol.39: 70-80; and Bingle, W.H. and Smit, J. 1994. "Alkaline Phosphatase and a Cellulase Reporter Protein Are Not Exported From the Cytoplasm When Fused to Large N-terminal Portions of the Caulobacter crescentus Surface (S)-Layer Protein". Can.J. Microbiol. 40:777-782.) and, cellulase (eg. by expression of the CenA gene of Cellulomonas fimi; see: Bingle, W.H. et al. (1993) [supra]; and Bingle, W.H. and Smit, J. (1994) [supra]).

Another advantageous application of this invention is the production of organisms that secrete and optionally present vaccine-candidate epitopes. For example, modified Caulobacter may be readily cultured in outdoor freshwater environments and would be particularly useful in fish vaccines. The two-dimensional crystalline array of the S-

protein layer of Caulobacter, which has a geometrically regular, repetitive structure, provides an ideal means for dense packing and presentation of a foreign epitope to an immune system in cases where the epitope is part of an intact S-layer in the bacterial cell surface.

This invention also provides an efficient expression system for polypeptides that may be harvested in large quantities relatively free of contaminants and protein of Caulobacter origin. Expression of a heterologous polypeptide fused with sufficient C-terminal amino acids of the S-layer protein to promote secretion of the heterologous polypeptide results in the accumulation of large quantities of secreted protein in the cell medium. In such cases, the chimeric protein does not have to be released from the cell surface. Furthermore, adjustment of the size of the S-layer protein portion can dictate whether the secreted chimeric protein is soluble or will precipitate in the cell medium. This embodiment may also be useful in cases where the Caulobacter is to express a foreign antigenic component and it is desired to minimize the amount of Caulobacter protein that is associated with the foreign antigen secreted by the Caulobacter.

Example 1: Production of Permissive Insertion Sites in C.crescentus

Using the restriction enzyme TaqI, a partial digestion of the rsaA gene in pTZ18U:rsaAΔP produced a group of linearized segments with random TaqI sites cleaved. The linearized segments were modified by use of the tagged linker mutagenesis procedure of Bingle and Smit (1991) [supra], using the 12-bp BamHI linker carried in plasmid pUC102K discussed in the general procedure above. Those products that produced a full-length protein in E. coli were ultimately transferred to pWBI (a minor variation of pWB9 that is replicated by Caulobacter), as described in the general procedure. The resulting construction was

introduced into a C. crescentus strain. Distinguishable events were retrieved and analyzed for the ability to produce a full-length protein in C. crescentus and to produce the crystalline S-layer on their surface and the approximate location of the insertion. Cells were screened for the presence of a S-layer protein of approximately 100kDa that is extracted from the surface of whole cells by 100 mM HEPES at ph2. The results of this screening together with the approximate positions of five successful events (and subsequently determined exact or specific insertion positions) are illustrated in Figure 5.

The above-described five positive events represent cases where the 4-amino acid insertion is tolerated with no effect on the S-layer function. The S-layers of the modified Caulobacter were indistinguishable from a wild-type S-layer. Thus, they have a higher potential for tolerating the addition of more foreign peptide material than less characterized sites. By producing 3 versions of the gene of interest, representing each possible reading frame (using standard linker addition technology), one may test each of these sites for suitability in expressing the desired activity. Also, by using restriction enzymes other than TaqI (such as AciI, HinPI or MspI) a larger library of BamHI insertions may be created.

Example 2: Insertion of Cadmium binding polypeptides Into Specific Sites

An insertion of the above described 12 bp linker was made at the TaqI site that corresponds to amino acid #188, frame #3 (see Figure 6; SEQ ID NO:6; and, SEQ ID NO:7). This created a unique BamHI site at that position. Because the precise position of the TaqI site could be assessed from the DNA sequence information available for the rsaA gene, the necessary translation frame was known and thus a single construction of a metallothionein gene was made.

This was done by excision of the coding sequence of monkey metallothionein II peptide (60 amino acids comprising 10 cysteine residues and having a molecular weight of about 5000) at known restriction sites and adapting the gene ends with BamHI linkers with appropriate base pair spacers for the needed translation frame.

After insertion into the BamHI site created at position 188, frame 3, several clones were examined by determining whether they could bind elevated levels of cadmium by the assay described below. The assay was necessary because the segment had equal probability of being inserted backwards. One clone that gave positive results was examined by electron microscopy and the presence of a normal S-layer was confirmed. The plasmid in the clone that gave positive results was also examined by DNA sequencing analysis, sequencing across the junction between the position 188 site and the 5′ side of the metallothionein gene. The sequence data confirmed correct orientation.

The plasmid-containing clone and relevant control strains were examined for the ability to bind several metals known to be bound by native metallothionein. This was done by growing the strains of bacteria in the presence of the metals at a concentration of 5ug/ml. After extensive washing of the cells to remove unbound metal, the cells were ashed by treatment at 500$\circ$C and the residue was dissolved in dilute nitric acid and examined for metal content by atomic absorption spectroscopy. The results from one round of data collection is shown in Table 1. In the case of cadmium and copper, an elevated level of bound metal is noted in the metallothionein-expressing strains.

Table 1

Caulobacter	Metal Ion Tested (ug/metal/OD unit of cells				
	Copper			Cadmium	Zinc
	Trial	1	2		
CB15 (wild-type, S-layer [+])	1.79		1.0	0.71	4.15
CB15KSAC (S-layer negative strain)	2.18	1.33	1.07		4.08
CB15KSAC/p188.3 (contains S-layer with linker insert only)	2.01		1.30	11.1	3.66
CB15KSAC/p188.3MT (S-layer with Metallothionein inserted)	2.79		3.09	19.1	3.00

Example 3: Investigation of Other Permissive Sites in rsaA Gene

A library of 240 BamHI linker insertions was created using the procedures of Example 1. Of the 240 insertions, 45 target sites in the rsaA gene were made with TaqI. 34 of the latter insertions were discarded because the clones contained deletions of rsaA DNA as well as the linker insertions. The remaining 11 resulted in 5 non-permissive and the 6 permissive sites described in Example 1. The remaining 195 insertions in the library were made using the enzymes HinPI, AciI, and MspI to create target sites as outlined in Example 1. Of the latter 195 insertions, 49 permissive sites were located for a total of 55. Of those sites scored as non-permissive, some may have had deletions of rsaA DNA at the linker insertion site. One BamHI linker insertion at a TaqI site thought to be permissive was later found by nucleotide sequencing to be located outside the rsaA structural gene reducing the total number of permissive sites to 54 from 55.

Figure 7 illustrates the approximate location by restriction mapping of 54 permissive sites. The results show that sites that will accept 2-4 amino acids while

still allowing the protein to be made and assembled into an S-layer are scattered up and down the protein. Furthermore, there is an unexpectedly high proportion of sites at which such insertions do not prevent expression and assembly of the S-layer. The results indicate that approximately 25-50% of in-frame linker insertions will be tolerated by the S-layer protein and the Caulobacter and that diverse regions of the protein will tolerate insertions. Thus, Caulobacter is an ideal candidate for expression of polypeptides fused with the S-layer and the presence of multiple permissive sites extending along the rsaA gene will permit the insertion of a plurality of the same or different peptides into the same RsaA protein molecule and expressed on the surface of a single Caulobacter.

Example 4: Further Studies with Cadmium binding Polypeptides

The results described for Example 3 indicated that it would be possible to insert metallothionein at multiple places in the RsaA protein and thereby enhance the metal binding capacity of such a transformed Caulobacter. However, when the procedures of Example 2 were repeated to insert the metallothionein coding sequence into others of the 54 permissive sites identified in the preceding Example in each case, the transformed Caulobacter did not secrete a chimeric protein and did not synthesize an S-layer. Furthermore, the transformed Caulobacter of Example 2 was stable as long as the transformants were frozen immediately after isolation. When continuously cultured for approximately one week, the transformants deleted the metallothionein portion of the S-layer and the S-layer protein returns to its normal size.

Consideration of the predicted amino acid sequence of the rsaA protein shows that the latter protein lacks cysteine residues whereas metallothionein has a high cysteine

content. It thus appeared that for secretion and long term expression of a RsaA chimeric protein, the heterologous polypeptide portions of the chimeric protein should not have high cysteine content and preferably, not be capable of forming multiple disulphide bonds in the chimeric protein in an aerobic environment.

Following the foregoing procedures, single and multiple copies of DNA encoding the synthetic cadmium binding peptide shown in Figure 10 (SEQ ID NO:11) was synthesized, inserted at the amino acid 277 site of <u>rsa</u>A using the above described Carrier cassette and was expressed in <u>C. crescentus</u>. The peptide has a single cysteine residue. Mild acid extracts of whole cells expressing the modified <u>rsa</u>A gene were subjected to SDS-PAGE for identification of S-layer proteins. The S-layer protein was expressed and secreted when there was from 1 to 3 copies of the cadmium binding peptide present at RsaA amino acid position 277. Insertion of 4 or more copies resulted in a dramatic reduction of S-layer protein released from the whole cells by mild acid treatment to barely detectable levels. Detection by autoradiography of RsaA protein <u>in vivo</u> labelled with 35 S-cysteine and <u>in vitro</u> with 125 I-iodoacetamide confirmed that the cadmium binding peptide was part of the chimeric RsaA protein. This demonstrates that <u>Caulobacter crescentus</u> is capable of secretion of a chimeric <u>rsa</u>A protein having a limited cysteine content and a limited capacity for disulphide bond formation within the chimeric protein.

<u>Example 5: Expression and Presentation of Antigenic Epitopes on Caulobacter Cell Surface</u>

Using the library of the 49 permissive sites other than those made with <u>Taq</u>I described in Example 3, the coding sequence for the 12-amino acid pilus peptide epitope lacking cysteine residues from <u>Pseudomonas aeruginosa</u> PAK pilin was inserted at the sites using the procedures

described above and employing the Carrier cassette shown in Figure 1. Positioning of the added DNA between the first Bam HI site and the Bgl II site permitted use of the latter site for making repeated insertions of DNA. The coding sequence for the peptide shown in Figure 8, including both cysteine residues was also inserted in separate experiments.

DNA coding for the peptide shown in Figure 8 (SEQ ID NO:8) was prepared by oligonucleotide synthesis of two anti-complementary strands. The transformed bacteria were screened for both production and presentation of the epitopes by the transformed Caulobacter by using standard Western immunoblot analysis (see: Burnette, W. N. 1981. "Western Blotting; Electrophoretic Transfer of Protein from Sodium Dodecyl-Polyacrylamide Gels to Unmodified Nitrocellulose and Radiographic Detection Antibody and Radioiodinated Protein A". Analytical Biochemistry 112:195-203) and by colony immunoblot tests in which the cells were not disrupted (see: Engleberg, N.C., et al. 1984. "Cloning an Expression of Legionella pneumophilia Antigens in Escherichia coli". Infection and Immunity 44:222-227). Anti-pilus monoclonal antibody obtained from Dr. Irvin, Dept. of Microbiology, University of Alberta (Canada) was used in the immunoblot analyses to detect the presence of the pilus epitope insert. The antibody (called PK99H) was prepared using purified Pseudomonas aeruginosa PAK pilin as the antigen and the monoclonal antibody against the 12 amino acid epitope was isolated by standard techniques using BALB/C mice as a source of ascites fluid. Reaction with the antibody in the whole cell colony immunoblot assay shows that the epitope is not only expressed in the transformed Caulobacter but is exposed on the S-layer surface overlying the cell in such a way that the epitope is available to the antibody. When the two cysteine residues of the pilin epitope were incorporated in the

chimeric protein, the protein was still expressed and secreted at normal levels.

Of the organisms screened, insertions of the pilus epitope at the following sites in the rsaA gene as determined by nucleotide sequencing resulted in a positive reaction with the antibody in the whole cell Colony immunoblot analysis: 69, 277, 353, 450, 485, 467, 551, 574, 622, 690, 723, and 944. The results show that the permissive sites that will accept polypeptides of the size of the pilus epitope are numerous and scattered across the rsaA gene.

Further studies with the pilus peptide resulted in successful expression and secretion of RsaA chimeric proteins have single copies of the peptide at the locations shown in Figure 11. Also, four and seven copies of the 12 amino acid pilus peptide were expressed and secreted as a RsaA chimeric protein when inserted at amino acids 277 and 551 respectively of the RsaA protein. However, insertions of the pilus peptide at amino acids 69, 277, 450, 551 and 622 resulted in a chimeric protein that did not attach to the cell surface and was released into the culture medium.

Example 6: Insertion of Large Polypeptides

Bacterial surface proteins from organisms other than *Caulobacter* described in the prior art are generally not known to accept polypeptides larger than about 60 amino acids within the structure of the surface protein. The procedures of the preceding Example were carried out in order to insert the coding sequence of a 109 amino acid epitope from IHNV virus coat glycoprotein at insertion sites identified in the preceding Example. The IHNV epitope was prepared by PCR and had the portion of the sequence shown in Figure 9 (SEQ ID NO:9) which is equivalent to amino acid residues 336-444 of the IHNV sequence described in: Koener, J.F. *et al*. 1987. "Nucleotide Sequence of a cDNA Clone Carrying the

Glycoprotein Gene of Infectious Hematopoietic Necrosis Virus, a Fish Rhabdovirus". Journal of Virology 61:1342-1349. Anti-IHNV polyclonal antibody against whole IHNV obtained from Dr. Joann Leong, Dept. of Microbiology, Oregon State University, U.S.A. (see: Xu, L. et al. 1991. "Epitope Mapping and Characterization of the Infectious Hematopoietic Necrosis Virus Glycoprotein, Using Fusion Proteins Synthesized in Escherichia coli". Journal of Virology 65:1611-1615) was used in the immunoblot assays described in the preceding Example to screen for Caulobacter that express and present the IHNV sequence on the surface of the S-layer of the Caulobacter. Reaction in the whole cell colony immunoblot assay was positive in respect of insertions at sites 450 and 551, and negative at a site which was at approximately amino acid 585.

The IHNV insert contains a single cysteine residue and is an extremely large insert for successful expression as a fusion product with a bacterial surface protein.

In further studies, the same 109 amino acid portion of the IHNV glycoprotein was inserted at amino acid 450 of the RsaA protein. The chimeric protein expressed and secreted by Caulobacter crescentus and was recovered from the cell culture medium. SDS-PAGE analysis of the recovered proteins showed that some of the chimeric proteins were smaller than the predicted rsaA chimeric protein but still bound anti-IHNV antibody. Analysis of these proteolytic products showed that cleavage of the chimeric protein occurred at an Arg residue encoded by the gene transfer cassette shown in Figure 1. Thus in some cases, adjustment of the nucleotide sequence at the interface of the polypeptide and rsaA coding sequences may be necessary to prevent expression of an arginine residue.

Example 7

Methods are described above for the insertion of 12-bp BamHI linker sites into a promoterless version of the rsaA gene. Because linker insertions involve the insertion of 12 bp (i.e. a multiple of three) an in-frame linker insertion resulted in every case. These linker sites are introduced to allow subsequent insertion of DNA encoding foreign peptide/proteins. Expression of such chimeric genes leads to the production of an entire full-length RsaA protein carrying the inserted heterologous amino acid sequence of interest. A number of BamHI site positions were identified above precisely by nucleotide sequencing. Four of the sites in the rsaA gene correspond to amino acid positions 188, 782, 905, 944 in the RsaA protein. For this example, an additional linker insertion was created at amino acid position 95 of the native gene (i.e. this gene carried its own promoter) using the same methodology. All five in-frame BamHI linker insertion sites were inserted in the rsaA so that the nucleotides of the linker DNA were read in the reading frame GGA/TCC (Figure 12).

Because all BamHI linker nucleotides were read in the same reading frame, the 5′ region of one rsaA gene carrying a BamHI linker insertion at one position could be combined with the 3′ region of an rsaA gene carrying another of the BamHI linker insertions to create in-frame deletions with a BamHI site at the joint between adjacent regions of rsaA. Using such a method, in-frame deletions of rsaA (ΔAA95-782) and rsaA(ΔAA188-782) were created.

DNA fragments encoding various C-terminal portions of the 1026 amino acid RsaA protein were isolated using the newly inserted BamHI linker sites as the 5′-terminus of the fragment and a HindIII site as the 3′ terminus of the fragment. These BamHI fragments were transferred to the BamHI/HindIII sites of pUC8 (J. Vieira, and J. Messing. 1982." The pUC Plasmids, an M13mp7-Derived System for Insertion Mutagenesis and Sequencing With Synthetic Universal Primers" Gene 19:259-268) creating "rsaA C-

terminal Segment Carrier plasmids" (Figure 12). The insertion into pUC8 also resulted in the creation of an in-frame fusion between the first 10 N-terminal amino acids of LacZa and the various C-terminal fragments (AA782-1026, AA905-1026 or AA944-1026) of RsaA. These LacZa:rsaA fusion proteins can be produced in C. crescentus using the lacZa transcription/translation initiation signals when introduced on appropriate plasmid vectors or direct insertion into the chromosome (see: W.H. Bingle, et al. 1993. "An All-Purpose Cellulase Reporter for Gene Fusion Studies and Application to the Paracrystalline Surface (S)-Layer Protein of Caulobacter crescentus." Can. J. Microbiol. 39:70-80).

Both types of constructions (i.e., the deletion versions and the C-terminal only segments) result in the production of proteins that are secreted in Caulobacter strains as highly modified RsaA proteins. The gene segments can also facilitate the secretion of heterologous polypeptides by insertion or fusion of appropriate DNA sequences at the unique BamH1 site that exists in each of the constructions. The following describes specific methods for doing so to create chimeric proteins capable of secretion in C. crescentus.

A- Creating fusions of desired sequences with C-terminal portions of rsaA -Method 1

The general process is as follows:

1) Inserting the desired sequence into the Carrier cassette. The following describes the specific manner in which heterologous sequences may be introduced into the Carrier cassette of Figure 1.

a) Insertion of a single copy of the desired gene segment.

Depending upon the length of the gene segment, two methods of construction may be used. For segments of up to about 30 amino acids, two oligonucleotides of appropriate sequence are chemically synthesized, annealed by mixing, heating and slow cooling and then ligated into the Carrier cassette. The oligonucleotides will also contain additional base pairs that recreate "sticky ends" of appropriate restriction endonuclease sites at each end of the duplex DNA that results from the annealing process.

For longer segments, PCR is used to amplify a region of a target DNA sequence. Oligonucleotides are synthesized that have sequence complementary to the boundaries of the desired sequence and which contain additional base pairs that recreate a "sticky end" of an appropriate restriction endonuclease site. In the present example oligonucleotides are made to produce products with the appropriate restriction endonuclease site for directional cloning into the Carrier cassette. PCR amplification of the desired sequence is then done by standard methods.

For both methods, the sticky ends prepared must be appropriate for an XhoI site at the 5′ terminus of the desired DNA sequence and StuI or SalI sites at the 3′ terminus; this places the desired gene segment in the correct orientation within the Carrier cassette. Reading frame continuity is maintained by appropriate design of the oligonucleotides used for the PCR step.

b) Preparation of multiple copies of the desired gene segment.

The Carrier cassette also allows production of multiple insert copies. A BglII site in the cassette is restored after removal of the promoterless antibiotic resistance gene; that site can be used to insert an additional copy of the Carrier/desired sequence insertion, using the terminal BamHI sites, because the "sticky ends"

produced by both BamHI and BglII are the same. This "piggy-back" insertion still maintains the correct reading frame throughout the construction. Any number of additional cycles of "piggy-backing" can be done because the BamHI/BglII ligation results in sequence which is no longer a substrate for either enzyme. The result is the production of cassettes of multiple copies of the desired sequence which can be transferred to appropriately modified S-layer protein genes with the same ease as a single copy. An additional feature of this method is that different heterologous sequences can be paired together in this multiple copy cassette with the same ease as multiple copies of the same heterologous sequence.

Example 7a. Insertion of an 109 amino acid segment of the IHNV surface glycoprotein to Carrier cassette.

Using the methods described, a PCR product was made that contained the DNA coding for amino acids 336 to 444 (Figure 9) of the major surface glycoprotein of the Infectious Hematopoietic Necrosis Virus (IHNV), which infects Salmonid fish.

Example 7b. Insertion of an 184 amino acid segment of the IHNV surface glycoprotein to Carrier cassette.

Using the methods described a PCR product was made that contained the DNA coding for amino acids 270 to 453 of the IHNV glycoprotein segment shown in Figure 9..

Example 7c. Insertion of single and multiple copies and an epitope of the Pseudomonas aeruginosa PAK pilus gene to Carrier cassette.

Oligonucleotides were constructed to code for the pilus epitope described in Example 5, which corresponds to a sequence at the extreme C-terminus of the pilus protein. Using the methods outlined in part A(1)(b) of this Example, 3 tandem copies were prepared.

2) Transfer of Carrier cassette to the rsaA C-terminal Segment Carrier plasmids. The constructions described in examples 7a and 7b above are then transferred to the rsaA C-terminal Segment Carrier plasmids, described above, resulting in an in-frame fusion of: a) a very short section of the betagalactosidase protein (10 amino acids), b) the desired sequence flanked by 2-3 amino acids derived from Carrier cassette sequence and c) the appropriate rsaA C-terminal segment. In some cases, the first codon of the rsaA C-terminal segment is converted to a different codon as a result of the fusion. For example, while the rsaA C-terminal segment may have coded for amino acids 944-1026 of RsaA, the resulting chimeric protein may only have amino acids 945-1026 native to RsaA.

Example 7d. Fusion of Carrier/109 AA and 184 IHNV segments to C-terminal rsaA segment AA782-1026.

This was done using the Carrier cassettes described in Examples 7a and 7b above and the AA782-1026 rsaA C-terminal Segment Carrier plasmid described above.

Example 7e. Fusion of Carrier/109 AA and 184 AA IHNV segments to C-terminal rsaA segment AA905-1026.

This was done using the Carrier cassettes described in Examples 7a and 7b above and the AA905-1026 rsaA C-terminal Segment Carrier plasmid described above.

Example 7f. Fusion of Carrier/109 AA and 184 AA IHNV segments to C-terminal rsaA segment AA944-1026.

This was done using the Carrier cassettes described in Examples 7a and 7b above and the AA944-1026 rsaA C-terminal Segment Carrier plasmid described above.

Example 7g. Fusion of Carrier/3x Pilus Epitope segment to C-terminal rsaA segment AA782-1026.

This was done using the Carrier cassettes described in Example 7c above and the AA782-1026 rsaA C-terminal Segment Carrier plasmid described above.

3) Expression of the desired fusion in an appropriate Caulobacter host strain.

a) Plasmid-based expression.

To create plasmid vectors that can be introduced and maintained in appropriate Caulobacter strains, the entire rsaA C-terminal Segment Carrier plasmids were fused to broad host range vectors pKT215 or pKT210 (see: M. Bagdasarian, et al. 1981." Specific-Purpose Cloning Vectors. II. Broad-Host-Range, High Copy Number RSF1010-Derived Vectors, and a Host-Vector System for Gene Cloning in Pseudomonas." Gene 16:237-247) using the unique HindIII restriction site present in each plasmid. The resulting plasmid is introduced into Caulobacter by conjugation or electroporation methods and is maintained by appropriate antibiotic selection.

The fusions described in examples 7d-7g were expressed in Caulobacter. In each case expression and secretion of the chimeric S-layer protein was detected by Western immunoblot analysis of electrophoretic gels of the cell culture supermutant employing the monoclonal antibody for each of the polypeptide epitopes. The transporter signal for secretion from Caulobacter must be in the C-terminal region of amino acids 945-1026 of the S-layer protein as all chimeric proteins in the examples were secreted. Precipitation of the chimeric protein occurred with the use of rsaA segment AA782-1026 but not AA944-1026. Recovery of precipitate using AA905-1026 was reduced as compared to AA782-1026.

b) Selection of appropriate Caulobacter host strains

In nearly all cases the use of a S-layer negative host strain is appropriate. C. crescentus strain CB2A and strain CB15aKSac fulfil this requirement. If it is important to ensure that all fusion protein is no longer attached to the cell surface, the use C. crescentus strains CB15Ca5KSac or CB15Ca10KSac are appropriate. These strains have additional mutations that result in the loss of the production of a specific species of surface lipopolysaccharide that has been demonstrated to be involved with the surface attachment of native S-layer protein as a 2-dimensional crystalline array (see: Walker S.G. et al 1994. "Characterization of Mutants of C. crescentus Defective in Surface Attachment of the Paracrystalline Surface Layer". J. Bacteriol. 176:6312-6323). Most often with the highly modified versions of the S-layer gene, this precaution is not necessary since virtually all regions of the gene that may have a role in the attachment process have been removed.

There are two types of growth media well suited to both propagation of Caulobacter for general purposes, including cloning steps, and also to produce the secreted and aggregated chimeric proteins. Example of the two types are: 1) PYE medium, a peptone and yeast extract based medium described in Walker et al, (1994) [supra], and 2) M6HiGG medium, a defined medium described in: Smit, J., et al 1981. "Caulobacter crescentus Pilin: Purification, Chemical Characterization and Amino-Terminal Amino Acid Sequence of a Structural Protein Regulated During Development". J. Biol. Chem. 256, 3092-3097. The latter medium is especially appropriate for preparation of the aggregated chimeric proteins since it permits growth to higher densities (therefore maximizing protein yield) and results in purer aggregated proteins since there are no

medium derived proteins to contaminate the chimeric proteins retrieved.

B- Creating Fusions of desired sequences with C-terminal portions of rsaA -Method 2.

Methods other than the use of the Carrier cassette plasmids are possible to create heterologous insertions into deletion versions of a S-layer gene or to create fusions with C-terminal portions of the S-layer protein. PCR may be used although other known methods may also be used. The general procedure is as follows:

1) Use PCR to prepare appropriate segments:

a) Preparation of amplified segment with appropriate ends is carried out in a manner similar to that described part A(1)(a) above. Oligonucleotides are designed and synthesized such that they will anneal to appropriate regions of the desired heterologous DNA and also contain "sticky ends" of appropriate sequence and frame so that the resulting PCR product can be directed inserted into appropriate modified S-layer genes.

b) Transfer to appropriate C-terminal *rsa*A segments is carried out by inserting the PCR products into the C-terminal segments AA782-1026, AA905-1026, or AA944-1026, as described in Examples 7d-7g above. In addition to the *Bam*HI site described, the *Eco*R1 restriction site could also be used as the 5' terminus of the incoming PCR segment, since this site is also available in the pUC8 vector and not in the S-layer gene, so long as the correct reading frame was maintained when designing the oligonucleotides used to prepare the PCR product.

2) Expression of the desired fusion in an appropriate *Caulobacter* host strain is carried out using the procedures outlined in part A(3) above.

C- Creating insertions of desired sequences into versions of a S-layer gene having large internal in-frame deletions.

The general process is as follows:

1) Creating appropriate in-frame deletions.

rsaA (ΔAA95-782) and rsaA(ΔAA188-782) were prepared as described above. Because most of the BamH1 linker insertion sites are in the same reading frame with respect to each other, it is possible to combine other pairs of 5' and 3' segments using the same general method, with the same result of maintenance of correct reading frame throughout. These deletion versions must then be tested individually to ensure that S-layer protein is still secreted by the Caulobacter.

2) Insertion of a Gene Segment Carrier cassette containing the desired sequences: as described at part A(1) above, carried out using the procedure described in part A(2) above.

Example 7h. Insertion of the 109 AA IHNV segment into rsaA (ΔAA95-782) and insertion of the 109 AA IHNV segment into rsaA(ΔAA188-782) is carried out as in Examples 7d-7g above. Expression of the desired genetic construction in appropriate C. crescentus strains is done using the procedures outlined in part A(3) above.

3) Alternate PCR procedures: can be used to prepare a heterologous segment for direct insertion into the BamHI site with the deletion versions of the rsaA gene. The procedure is essentially the same as described in part B(1) above.

<u>Example 8. (Transfer to the native S-layer gene chromosomal site as a single crossover event).</u>

The fusion of the Carrier cassette with appropriate heterologous DNA segments to a C-terminal S-layer protein segment plasmid results in a pUC8-based plasmid that is not maintained in <u>Caulobacter</u>. Selection for the antibiotic marker on the plasmid results in detection of the rescue events. Most commonly these are single crossover homologous recombination events. The result is a direct insertion of the entire plasmid into the chromosome. Thus the resident copy of the S-layer gene remains unchanged as well as the incoming highly modified S-layer gene. In such cases it may be desirable to use <u>Caulobacter</u> strains in which the resident S-layer gene is inactivated in known ways. One example is the use of <u>C</u>. <u>crescentus</u> strain CB15AKSac; this strain has an antibiotic resistance gene cassette introduced at a position in the S-layer gene about 25% of the way from the 5′ terminus.

<u>Example 9. (Transfer to the native S-layer gene chromosomal site as a double crossover event).</u>

In certain cases it may be desirable to completely exchange the resident S-layer gene copy with the incoming highly modified version. One method is the incorporation of a <u>sac</u>B gene cassette (Hynes, M.F., <u>et al</u>. 1989. "Direct Selection for Curing and Deletion of Rhizobium Plasmids Using Transposons Carrying the <u>Bacillus</u> <u>subtilis</u> sacB Gene." Gene 78: 111-119) into the pUC8 based plasmids carrying the desired chimeric gene construction. This cassette contains a levansucrase gene from <u>Bacillus</u> <u>subtilis</u> that, in the presence of sucrose, is thought to result in the production of a sugar polymer that is toxic to most bacteria when expressed inside the cell. One first selects for the single crossover event as described in Example 8. Subsequent growth on sucrose-containing medium results in the death of all cells except those that lose the offending <u>sac</u>B gene by homologous recombination within

the 2 adjacent gene copies. Two events are possible; restoration of the resident copy of the S-layer gene or replacement of the resident copy with the incoming modified gene (the latter is the desired event). A screen with insertion DNA as probe or antibody specific to the heterologous gene product identifies successful gene replacement events. The method requires that the S-layer gene sequence or native sequence immediately adjacent to the S-layer gene be on both sides of the heterologous sequence (ie, Carrier cassette sequence plus heterologous DNA) and in the present case is best suited for the deletion versions of the S-layer gene.

Other methods are available for the delivery of genes to the chromosome of a Caulobacter. Methods involving the use of the transposons Tn5 and Tn7 as a means of delivery of genes to random chromosome locations are available (see: Barry, G.F. 1988 "A Broad-Host-Range Shuttle System for Gene Insertion into the Chromosomes of Gram-Negative Bacteria." Gene 71:75-84.). The use of the xylose utilization operon as a target for chromosome insertion have also been described. This method involves the incorporation of a portion that operon into the pUC8 based plasmid constructions described above. This allows homologous recombination within the xylose operon as a means of plasmid rescue. Loss of the the ability to use xylose as a nutrition source is used as the means of confirming the rescue event.

This invention now being described, it will be apparent to one of ordinary skill in the art that changes and modifications can be made thereto without departing from the spirit or scope of the appended claims.

We Claim:

1. A DNA construct comprising one or more restriction sites for facilitating insertion of DNA to the construct upstream of DNA encoding a C-terminal region of a Caulobacter S-layer protein comprising at least the 82 C-terminal amino acids of the S-layer protein.

2. A DNA construct comprising DNA encoding a heterologous polypeptide sequence not present in a Caulobacter S-layer protein upstream from and in-frame with DNA encoding a C-terminal region of a Caulobacter S-layer protein comprising at least the 82 C-terminal amino acids of the S-layer protein.

3. The DNA construct of claim 1 or 2 further comprising an operably linked promoter recognized by Caulobacter.

4. The DNA construct of claim 1, 2 or 3 wherein said C-terminal region corresponds to amino acids 945-1026 of the RsaA protein of C. crescentus.

5. The DNA construct of claim 1, 2 or 3 wherein said C-terminal region comprises amino acids corresponding to about amino acids 850-1026 of the RsaA protein of C. crescentus.

6. The DNA construct of claim 1, 2 or 3 wherein said C-terminal region comprises amino acids corresponding to about amino acids 782-1026 of the RsaA protein of C. crescentus.

7. The DNA construct of claim 2, 3, 4, 5 or 6 wherein the heterologous polypeptide sequence is of one or more polypeptides of up to about 200 amino acids in length.

8. A bacterial cell comprising a DNA construct comprising DNA encoding a heterologous polypeptide not present in a S-

layer protein of Caulobacter upstream from and in-frame with DNA encoding a C-terminal region of a Caulobacter S-layer protein, wherein the C-terminal region comprises at least the 82 C-terminal amino acids of the S-layer protein.

9. The cell of claim 8 wherein the cell is a Caulobacter.

10. The cell of claim 9 wherein the DNA construct further comprises an operably linked promoter recognized by the Caulobacter and wherein the DNA construct is expressed in the cell and the protein so expressed is secreted by the cell.

11. The cell of claim 8, 9, or 10 wherein the C-terminal region corresponds to amino acids 945-1026 of the RsaA protein of C. crescentus.

12. The cell of claim 8, 9, or 10 wherein the C-terminal region comprises amino acids corresponding to about amino acids 850-1026 of the RsaA protein of C. crescentus.

13. The cell of claim 8, 9, or 10 wherein the C-terminal region comprises amino acids corresponding to about amino acids 782-1026 of the RsaA protein of C. crescentus.

14. The cell of claim 8, 9, 10, 11, 12, or 13 wherein the heterologous polypeptide is one or more polypeptides of up to about 200 amino acids in length.

15. The cell of claim 10 wherein the cell forms a S-layer comprising the heterologous polypeptide on a surface of the cell.

16. The cell of claim 8, 9, 10, 11, 12, 13, or 15 wherein the heterologous polypeptide sequence is one of more polypeptide of up to about 60 amino acids in length.

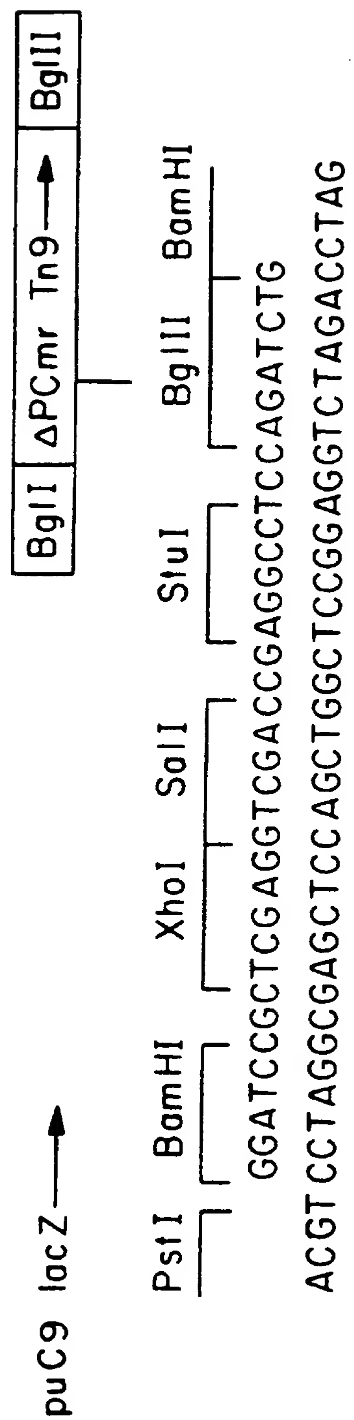

puC9 lacZ
PstI
BamHI
XhoI
SalI
StuI
BglII
BamHI
GGATCCGCTCGAGGTCGACCGAGGCCTCCAGATCTG
ACGTCCTAGGCGAGCTCCAGCTGGCTCCGGAGGTCTAGACCTAG
BglI
ΔPCmr Tn9
BglII

FIG. 1

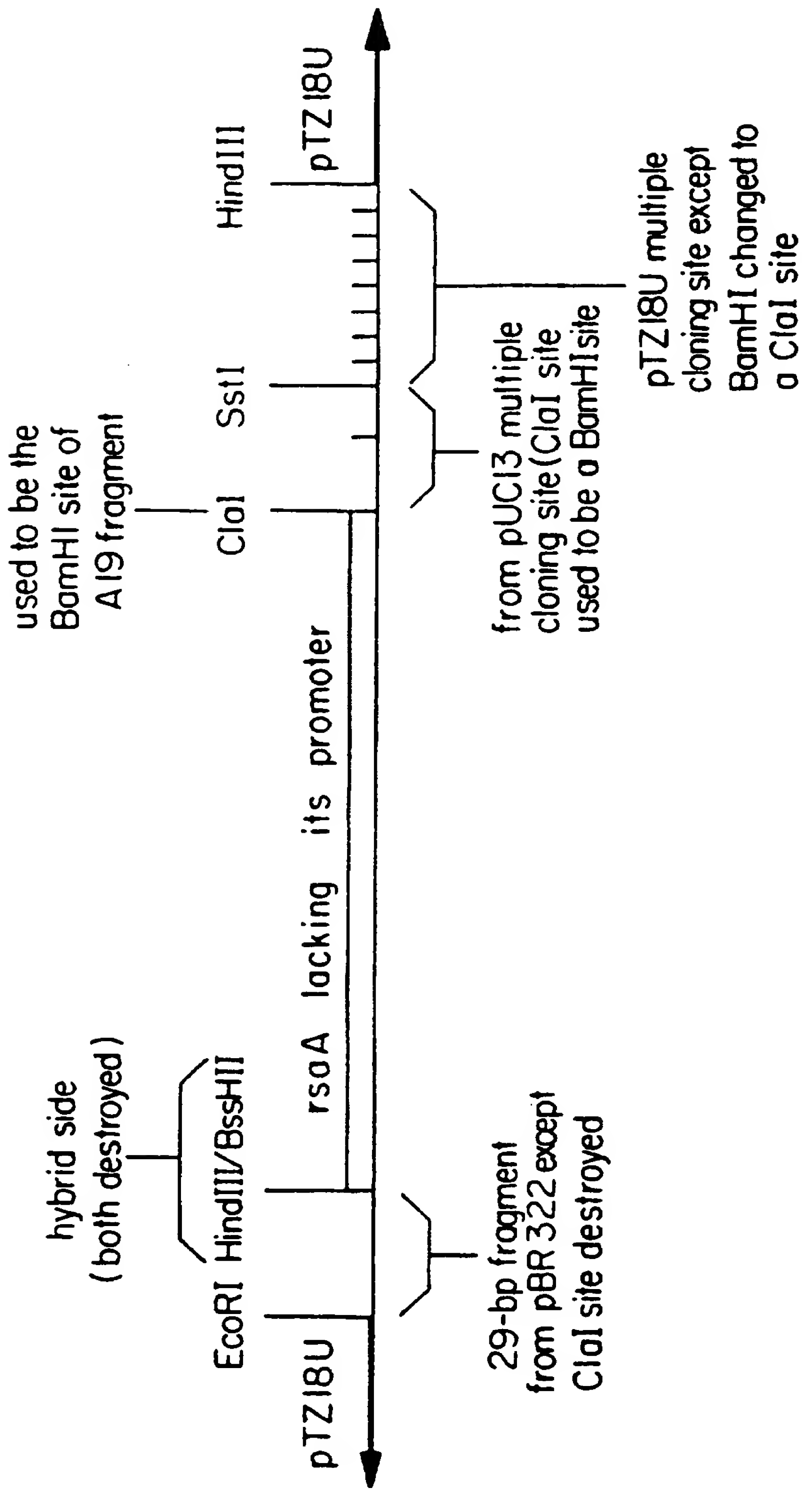

hybrid side
(both destroyed)
used to be the
BamHI site of
A19 fragment
pTZ18U
EcoRI
HindIII/BssHII
ClaI
SstI
HindIII
pTZ18U
rsaA lacking its promoter
29-bp fragment
from pBR 322 except
ClaI site destroyed
from pUC13 multiple
cloning site (ClaI site
used to be a BamHI site
pTZ18U multiple
cloning site except
BamHI changed to
a ClaI site

FIG. 2

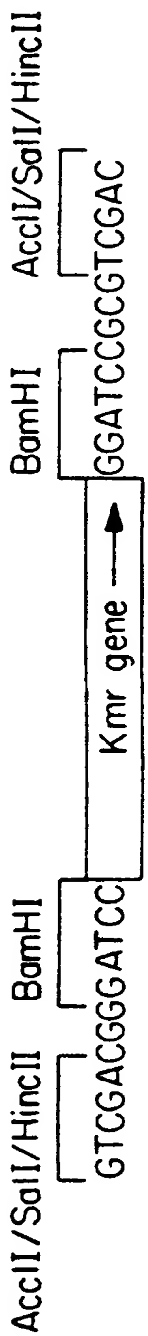

AccII/SalI/HincII
GTCGACGGGATCC
BamHI
Kmr gene
BamHI
GGATCCGCGTCGAC
AccI/SalI/HincII

FIG. 3

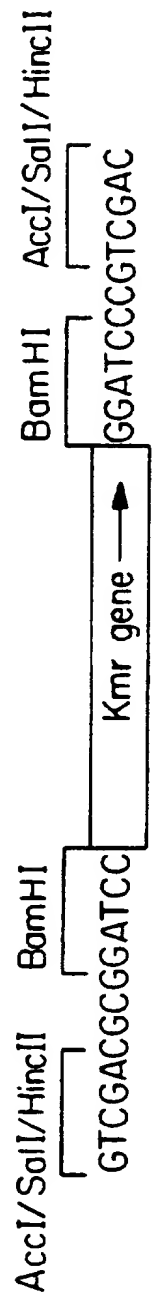

AccI/SalI/HincII
GTCGACGCGGATCC
BamHI
Kmr gene
BamHI
GGATCCCGTCGAC
AccI/SalI/HincII

FIG. 4

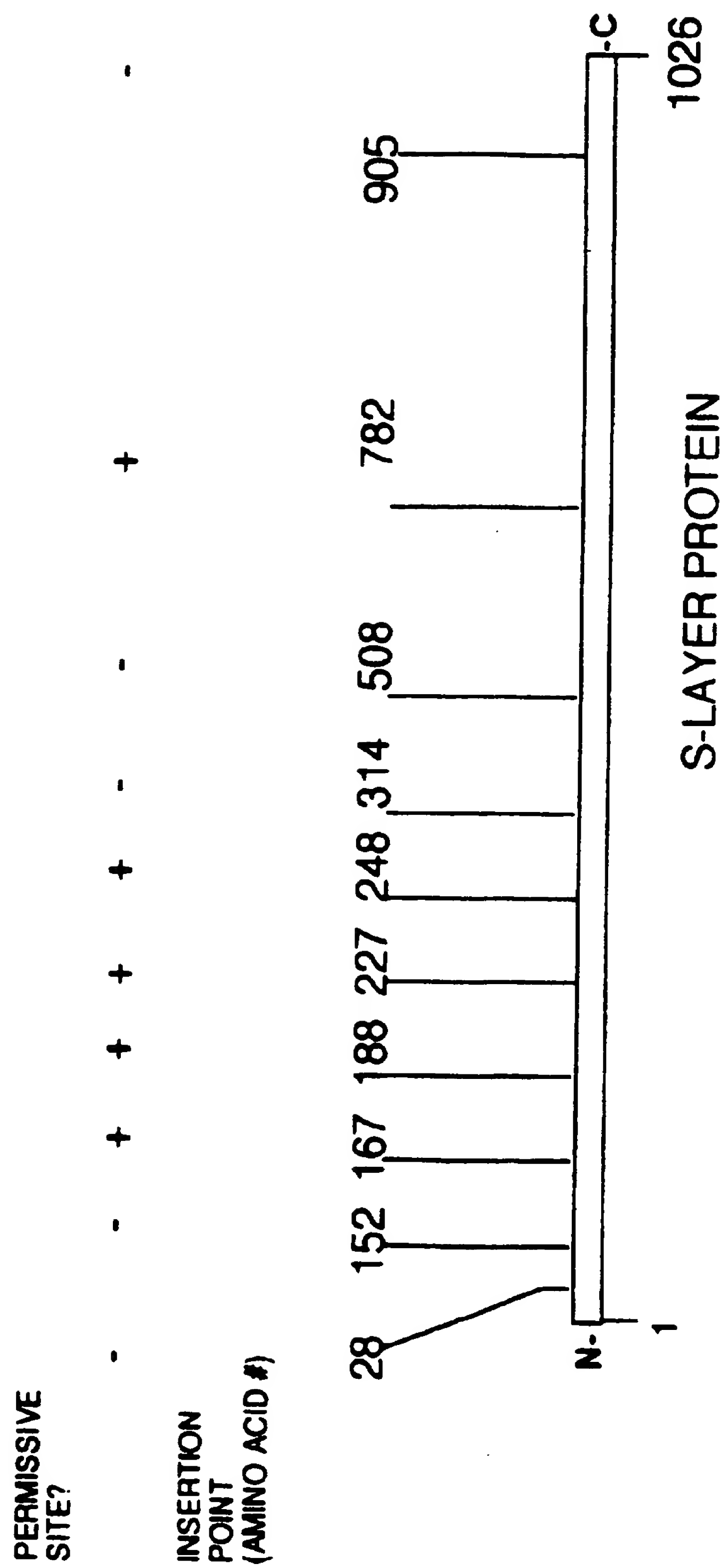

PERMISSIVE SITE?
INSERTION POINT (AMINO ACID #)
-
28
-
152
+
167
+
188
+
227
+
248
-
314
-
508
+
782
-
905
N-
1
-C
1026
S-LAYER PROTEIN

FIG. 5

```
GCTATTGTCGACGTATGACGTTTGCTCTATAGCCATCGCTGCTCCCATGCGCGCCACTCGGTCGCAGGGGGTGTGGGATTTTTTTGGGAGACAATCCTC
     -35                     -10                                                             S.D.

  1  M A Y T T A Q L V T A Y T N A N L G K A P D A A T T L T L D A Y A T
  1  ATGGCCTATACGACGGCCCAGTTGGTGACTGCGTACACCAACGCCAACCTCGGCAAGGCGCCTGACGCCGCCACCACGCTGACGCTCGACGCGTACGCGA

 35  Q T Q T G G L S D A A A L T N T L K L V N S T T A V A I Q T Y Q F
100  CTCAAACCCAGACGGGCGGCCTCTCGGACGCCGCTGCGCTGACCAACACCCTGAAGCTGGTCAACAGCACGACGGCTGTTGCCATCCAGACCTACCAGTT

 68  F T G V A P S A A G L D F L V D S T T N T N D L N D A Y Y S K F A
200  CTTCACCGGCGTTGCCCCGTCGGCCGCTGGTCTGGACTTCCTGGTCGACTCGACCACCAACACCAACGACCTGAACGACGCGTACTACTCGAAGTTCGCT

101  Q E N R F I N F S I N L A T G A G A G A T A F A A A Y T G V S Y A Q
300  CAGGAAAACCGCTTCATCAACTTCTCGATCAACCTGGCCACGGGCGCCGGCGCCGGCGCGACGGCTTTCGCCGCCGCCTACACGGGCGTTTCGTACGCCC

135  T V A T A Y D K I I G N A V A T A A G V D V A A A V A F L S R Q A
400  AGACGGTCGCCACCGCCTATGACAAGATCATCGGCAACGCCGTCGCGACCGCCGCTGGCGTCGACGTCGCGGCCGCCGTGGCTTTCCTGAGCCGCCAGGC

168  N I D Y L T A F V R A N T P F T A A A D I D L A V K A A L I G T I
500  CAACATCGACTACCTGACCGCCTTCGTGCGCGCCAACACGCCGTTCACGGCCGCTGCCGACATCGATCTGGCCGTCAAGGCCGCCCTGATCGGCACCATC

201  L N A A T V S G I G G Y A T A T A A M I N D L S D G A L S T D N A A
600  CTGAACGCCGCCACGGTGTCGGGCATCGGTGGTTACGCGACCGCCACGGCCGCGATGATCAACGACCTGTCGGACGGCGCCCTGTCGACCGACAACGCGG

235  G V N L F T A Y P S S G V S G S T L S L T T G T D T L T G T A N N
700  CTGGCGTGAACCTGTTCACCGCCTATCCGTCGTCGGGCGTGTCGGGTTCGACCCTCTCGCTGACCACCGGCACCGACACCCTGACGGGCACCGCCAACAA

268  D T F V A G E V A G A A T L T V G D T L S G G A G T D V L N W V Q
800  CGACACGTTCGTTGCGGGTGAAGTCGCCGGCGCTGCGACCCTGACCGTTGGCGACACCCTGAGCGGCGGTGCTGGCACCGACGTCCTGAACTGGGTGCAA

301  A A A V T A L P T G V T I S G I E T M N V T S G A A I T L N T S S G
900  GCTGCTGCGGTTACGGCTCTGCCGACCGGCGTGACGATCTCGGGCATCGAAACGATGAACGTGACGTCGGGCGCTGCGATCACCCTGAACACGTCTTCGG
```

FIG. 6a

6/11

```
335   V T G L T A L N T N T S G A A Q T V T A G A G Q N L T A T T A A Q
1000  GCGTGACGGGTCTGACCGCCCTGAACACCAACACCAGCGGCGCGGCTCAAACCGTCACCGCCGGCGCTGGCCAGAACCTGACCGCCACGACCGCCGCTCA

368   A A N N V A V D G R A N V T V A S T G V T S G T T T V G A N S A A
1100  AGCCGCGAACAACGTCGCCGTCGACGGGCGCGCCAACGTCACCGTCGCCTCGACGGGCGTGACCTCGGGCACGACCACGGTCGGCGCCAACTCGGCCGCT

401   S G T V S V S V A N S S T T T T G A I A V T G G T A V T V A Q T A G
1200  TCGGGCACCGTGTCGGTGAGCGTCGCGAACTCGAGCACGACCACCACGGGCGCTATCGCCGTGACCGGTGGTACGGCCGTGACCGTGGCTCAAACGGCCG

435   N A V N T T L T Q A D V T V T G N S S T T A V T V T Q T A A A T A
1300  GCAACGCCGTGAACACCACGTTGACGCAAGCCGACGTGACCGTGACCGGTAACTCCAGCACCACGGCCGTGACGGTCACCCAAACCGCCGCCGCCACCGC

468   G A T V A G R V N G A V T I T D S A A A S A T T A G K I A T V T L
1400  CGGCGCTACGGTCGCCGGTCGCGTCAACGGCGCTGTGACGATCACCGACTCTGCCGCCGCCTCGGCCACGACCGCCGGCAAGATCGCCACGGTCACCCTG

501   G S F G A A T I D S S A L T T V N L S G T G T S L G I G R G A L T A
1500  GGCAGCTTCGGCGCCGCCACGATCGACTCGAGCGCTCTGACGACCGTCAACCTGTCGGGCACGGGCACCTCGCTCGGCATCGGCCGCGGCGCTCTGACCG

535   T P T A N T L T L N V N G L T T T G A I T D S E A A A D D G F T T
1600  CCACGCCGACCGCCAACACCCTGACCCTGAACGTCAATGGTCTGACGACGACCGGCGCGATCACGGACTCGGAAGCGGCTGCTGACGATGGTTTCACCAC

568   I N I A G S T A S S T I A S L V A A D A T T L N I S G D A R V T I
1700  CATCAACATCGCTGGTTCGACCGCCTCTTCGACGATCGCCAGCCTGGTGGCCGCCGACGCGACGACCCTGAACATCTCGGGCGACGCTCGCGTCACGATC

601   T S H T A A A L T G I T V T N S V G A T L G A E L A T G L V F T G G
1800  ACCTCGCACACCGCTGCCGCCCTGACGGGCATCACGGTGACCAACAGCGTTGGTGCGACCCTCGGCGCCGAACTGGCGACCGGTCTGGTCTTCACGGGCG

635   A G R D S I L L G A T T K A I V M G A G D D T V T V S S A T L G A
1900  GCGCTGGCCGTGACTCGATCCTGCTGGGCGCCACGACCAAGGCGATCGTCATGGGCGCCGGCGACGACACCGTCACCGTCAGCTCGGCGACCCTGGGCGC

668   G G S V N G G D G T D V L V A N V N G S S F S A D P A F G G F E T
2000  TGGTGGTTCGGTCAACGGCGGCGACGGCACCGACGTTCTGGTGGCCAACGTCAACGGTTCGTCGTTCAGCGCTGACCCGGCCTTCGGCGGCTTCGAAACC

701   L R V A G A A A Q G S H N A N G F T A L Q L G A T A G A T T F T N V
2100  CTCCGCGTCGCTGGCGCGGCGGCTCAAGGCTCGCACAACGCCAACGGCTTCACGGCTCTGCAACTGGGCGCGACGGCGGGTGCGACGACCTTCACCAACG
```

FIG. 6b

```
735   A V N V G L T V L A A P T G T T T V T L A N A T G T S D V F N L T
2200  TTGCGGTGAATGTCGGCCTGACCGTTCTGGCGGCTCCGACCGGTACGACGACCGTGACCCTGGCCAACGCCACGGGCACCTCGGACGTGTTCAACCTGAC

768   L S S S A A L A A G T V A L A G V E T V N I A A T D T N T T A H V
2300  CCTGTCGTCCTCGGCCGCTCTGGCCGCTGGTACGGTTGCGCTGGCTGGCGTCGAGACGGTGAACATCGCCGCCACCGACACCAACACGACCGCTCACGTC

801   D T L T L Q A T S A K S I V V T G N A G L N L T N T G N T A V T S F
2400  GACACGCTGACGCTGCAAGCCACCTCGGCCAAGTCGATCGTGGTGACGGGCAACGCCGGTCTGAACCTGACCAACACCGGCAACACGGCTGTCACCAGCT

835   D A S A V T G T A P A V T F V S A N T T V G E V V I I R G G A G A
2500  TCGACGCCAGCGCCGTCACCGGCACGGCTCCGGCTGTGACCTTCGTGTCGGCCAACACCACGGTGGGTGAAGTCGTCACGATCCGCGGCGGCGCTGGCGC

      *         *           *           *
868   D S L I G S A I A N D I I I G G A G A D I L V Y I G G I D I F I G
2600  CGACTCGCTGACCGGTTCGGCCACCGCCAATGACACCATCATCGGTGGCGCTGGCGCTGACACCCTGGTCTACACCGGCGGTACGGACACCTTCACGGGT

              *
901   G I G A D I F D I N A I G T S T A F V T I T D A A V G D K L D L V G
2700  GGCACGGGCGCGGATATCTTCGATATCAACGCTATCGGCACCTCGACCGCTTTCGTGACGATCACCGACGCCGCTGTCGGCGACAAGCTCGACCTCGTCG

935   I S T N G A I A D G A F G A A V T L G A A A T L A Q Y L D A A A A
2800  GCATCTCGACGAACGGCGCTATCGCTGACGGCGCCTTCGGCGCTGCGGTCACCCTGGGCGCTGCTGCGACCCTGGCTCAGTACCTGGACGCTGCTGCTGC

968   G D G S G T S V A K W F Q F G G D T Y V V V D S S A G A T F V S G
2900  CGGCGACGGCAGCGGCACCTCGGTTGCCAAGTGGTTCCAGTTCGGCGGCGACACCTATGTCGTCGTTGACAGCTCGGCTGGCGCGACCTTCGTCAGCGGC

1001  A D A V I K L T G L V T L T T S A F A T E V L T L A end
3000  GCTGACGCGGTGATCAAGCTGACCGGTCTGGTCACGCTGACCACCTCGGCCTTCGCCACCGAAGTCCTGACGCTCGCCTAAGCGAACGTCTGATCCTCGC

3100  CTAGGCGAGGATCGCTAGACTAAGAGACCCCGTCTTCCGAAAGGGAGGCGGGGTCTTTCTTATGGGCGCTACGCGCTGGCCGGCCTTGCCTAGTTCCGGT
```

FIG. 6c

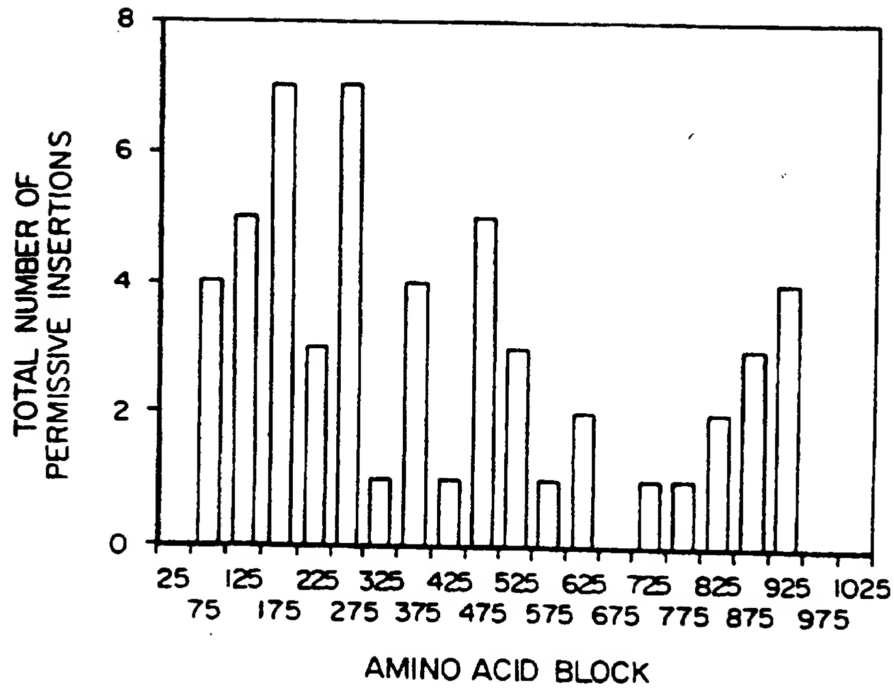

TOTAL NUMBER OF PERMISSIVE INSERTIONS
0
2
4
6
8
25 125 225 325 425 525 625 725 825 925 1025
75 175 275 375 475 575 675 775 875 975
AMINO ACID BLOCK

FIG. 7

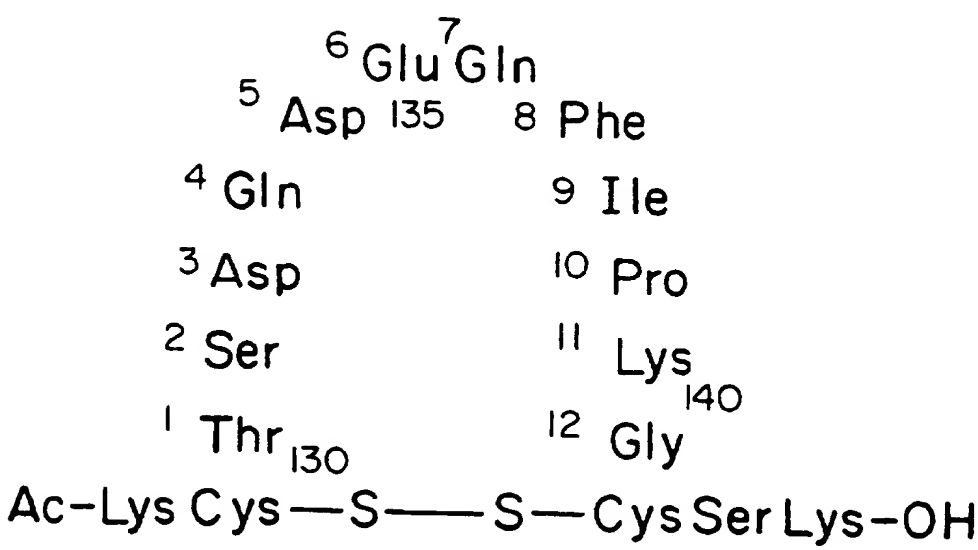

6 Glu 7 Gln
5 Asp 135 8 Phe
4 Gln 9 Ile
3 Asp 10 Pro
2 Ser 11 Lys
140
1 Thr 130 12 Gly
Ac-Lys Cys—S——S—Cys Ser Lys-OH

FIG. 8

1/1 31/11
GAA TAC AAT TCT GGA GCA GAA ATC CTC TCG TTC CCG AAG TGT GAG GAC AAG ACG ATG GGG
glu tyr asn ser gly ala glu ile leu ser phe pro lys cys glu asp lys thr met gly

61/21 91/31
ATG AGG GGA AAC TTG GAT GAC TTT GCC TAT CTA GAC GAT CTG GTG AAG GCC TCT GAG AGC
met arg gly asn leu asp asp phe ala tyr leu asp asp leu val lys ala ser glu ser

121/41 151/51
AGA GAG GAA TGT CTT GAG GCG CAC GCC GAG ATA ATA TCA ACA AAC AGT GTG ACT CCA TAC
arg glu glu cys leu glu ala his ala glu ile ile ser thr asn ser val thr pro tyr

181/61 211/71
CTC CTA TCC AAG TTC CGA TCT CCA CAT CCC GGA ATA AAT GAC GTC TAC GCT ATG CAC AAA
leu leu ser lys phe arg ser pro his pro gly ile asn asp val tyr ala met his lys

241/81 271/91
GGC TCC ATC TAT CAC GGG ATG TGC ATG ACG GTC GCT GTG GAC GAG GTA TCC AAG GAC AGG
gly ser ile tyr his gly met cys met thr val ala val asp glu val ser lys asp arg

301/101 331/111
ACG ACG TAC AGG GCC CAT CGC GCT ACC AGC TTC ACG AAA TGG GAA CGA CCC TTT GGG GAT
thr thr tyr arg ala his arg ala thr ser phe thr lys trp glu arg pro phe gly asp

361/121 391/131
GAG TGG GAG GGC TTT CAC GGA TTG CAC GGA AAC AAC ACC ACC ATT ATT CCA GAC CTG GAG
glu trp glu gly phe his gly leu his gly asn asn thr thr ile ile pro asp leu glu

421/141 451/151
AAA TAC GTC GCC CAG TAC AAG ACG AGC ATG ATG GAA CCG ATG AGC ATC AAA TCC GTA CCC
lys tyr val ala gln tyr lys thr ser met met glu pro met ser ile lys ser val pro

481/161 511/171
CAT CCA AGC ATC CTG GCC TTC TAC AAT GAG ACA GAC TTA TCA GGG ATC TCC ATC AGG AAA
his pro ser ile leu ala phe tyr asn glu thr asp leu ser gly ile ser ile arg lys

541/181
TTG GAC TCA TTC
leu asp ser phe

FIG. 9

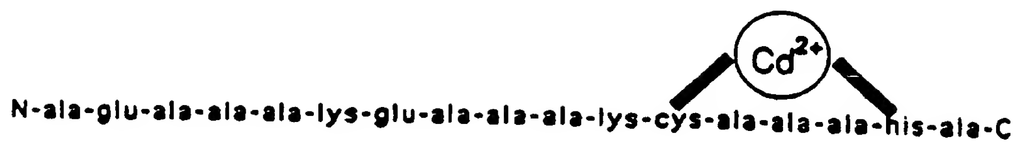

Cd2+
N-ala-glu-ala-ala-ala-lys-glu-ala-ala-ala-lys-cys-ala-ala-ala-his-ala-C

FIG. 10

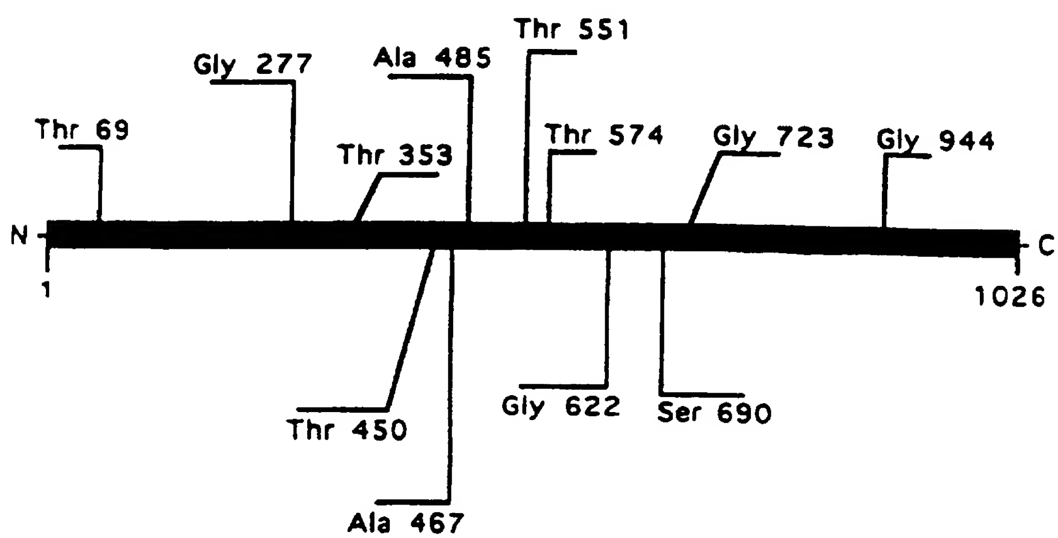

Thr 551
Gly 277
Ala 485
Thr 69
Thr 353
Thr 574
Gly 723
Gly 944
N
C
1
1026
Thr 450
Gly 622
Ser 690
Ala 467

FIG. 11

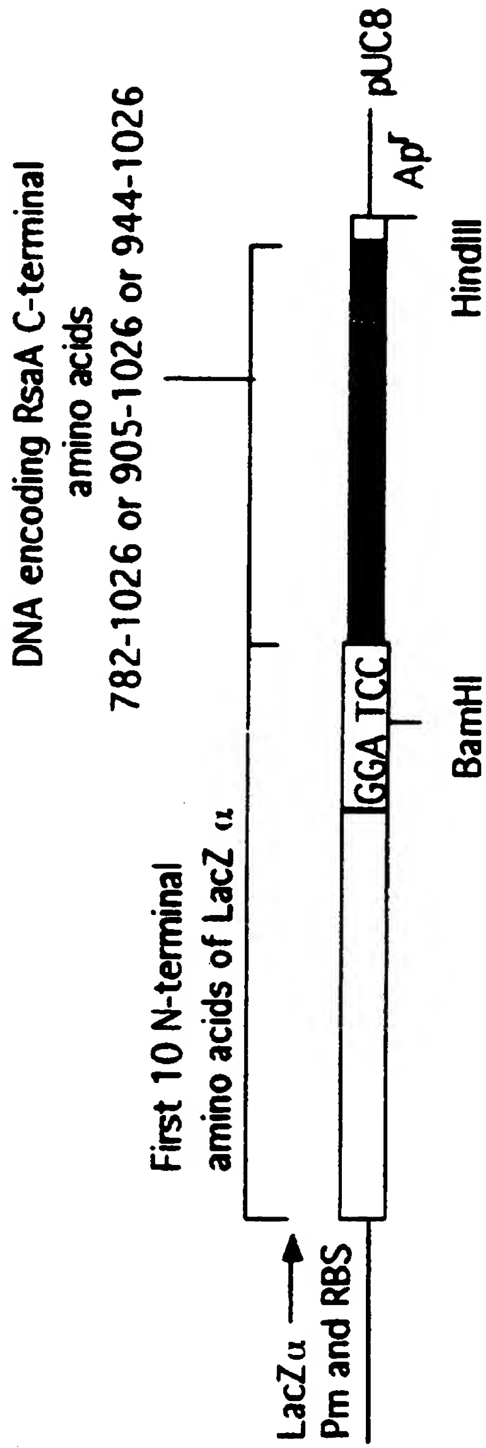

DNA encoding RsaA C-terminal
amino acids
782-1026 or 905-1026 or 944-1026
First 10 N-terminal
amino acids of LacZ α
LacZα
Pm and RBS
GGA TCC
BamHI
HindIII
Ap^r
pUC8

FIG. 12

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 97/00167

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/31 C12N15/47 C12N15/62 C12N15/74 C12N1/21
C07K14/145 C07K14/21

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ABSTRACTS OF THE GENERAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, vol. 95, 1995, page 525 XP002035105 NOMELLINI J. ET AL.: "Insertion of heterologous peptides within the surface-layer protein of Caulobacter crescentus" see the whole document ---	1-18,21

-/--

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

3

Date of the actual completion of the international search	Date of mailing of the international search report
11 July 1997	18.07.97

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Kania, T

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE WPI Section Ch, Week 9409 Derwent Publications Ltd., London, GB; Class C12,Page 027, AN 94-066249 XP002035014 "Production of heterologous polypeptides in bacteria, particularly Caulobacter- by expression of a fusion product of the polypeptide sequence and a bacterial S-layer protein gene" & CA 2 090 549 (UNIVERSITY OF BRITISH COLUMBIA) , 10 December 1993 see abstract	1-18
P,X	& US 5 500 353 A (UNIVERSITY OF BRITISH COLUMBIA) 19 March 1996 see the whole document ---	1-18,21
A	CANADIAN JOURNAL OF MICROBIOLOGY, vol. 40, no. 9, September 1994, pages 777-782, XP000674689 BINGLE W. AND SMIT J.: "Alkaline phosphatase and a cellulase reporter protein are not exported from the cytoplasm when fused to large N-terminal portions of the Caulobacter crescentus surface (S)-layer protein" cited in the application * see the whole document, esp. p.781,l.1-12 * ---	1-21
A	CANADIAN JOURNAL OF MICROBIOLOGY, vol. 38, no. 3, March 1992, pages 193-202, XP002034943 GILCHRIST A. ET AL.: "Nucleotide sequence of the gene encoding the Caulobacter crescentus paracrystalline surface layer protein" cited in the application see the whole document ---	1-21
A	JOURNAL OF VIROLOGY, vol. 61, 1987, pages 1342-1349, XP002034944 KOENER J. ET AL.: "Nucleotide sequence of a cDNA clone carrying the glycoprotein gene of infectious hematopoetic necrosis virus, a fish rhabdovirus" cited in the application see the whole document ---	21

3

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	JOURNAL OF BACTERIOLOGY, vol. 179, no. 3, February 1997, pages 601-611, XP002034945 BINGLE W. ET AL.: "Linker mutagenesis of the Caulobacter crescentus S-layer protein: toward a definition of an N-terminal anchoring region and a C-terminal secretion signal and the potential for heterologous protein secretion" see the whole document ---	1-10,13, 14, 16-18, 20,21
T	CANADIAN JOURNAL OF MICROBIOLOGY, vol. 42, no. 7, July 1996, pages 672-684, XP002034946 BINGLE W. ET AL.: "The extreme N-terminus of the Caulobacter crescentus surface-layer protein directs export of passenger proteins from the cytoplasm but is not required for secretion of the native protein" see the whole document -----	1-21

3